

1600

RAW SEQUENCE LISTING

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:44

Input Set : A:\ST96042Asq1t.txt

Output Set: N:\CRF3\11082001\I595947A.raw

3 <110> APPLICANT: ICARD-LIEPKALNS, Christine
 4 MALLET, Jacques
 5 RAVASSARD, Philippe
 7 <120> TITLE OF INVENTION: POLYPEPTIDES OF THE "BASIC-HELIX-LOOP-HELIX" bHLH
 8 FAMILY, CORRESPONDING NUCLEIC ACID SEQUENCES
 10 <130> FILE REFERENCE: ST96042A-US
 12 <140> CURRENT APPLICATION NUMBER: US 09/595,947A
 13 <141> CURRENT FILING DATE: 2000-06-16
 15 <150> PRIOR APPLICATION NUMBER: FR96/15651 *OK*
 16 <151> PRIOR FILING DATE: 1996-12-19 *OK*
 18 <150> PRIOR APPLICATION NUMBER: PCT/FR97/02368
 19 <151> PRIOR FILING DATE: 1997-12-19
 21 <160> NUMBER OF SEQ ID NOS: 28
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 1460
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Rattus norvegicus
 30 <400> SEQUENCE: 1
 31 gcaggtagcg agaggagcag tccctggggc cccgttgctg attggcccg ggcacaggca 60
 32 gcagcccggc aggcacgctc ctgggtccggg cagagcagat aaagcgtgcc aggggacaca 120
 33 cgattagcag ctccagaagtc cctctgggtc tcaccactgc acagaggccg aggaccccct 180
 34 ccgagcttct ttgtgcctc cagacgcaat ttactccagg cgaggggccc tgcagctcag 240
 35 caaaaacttcg aagcgagcag aggggttcag ctatccaccg ctgcttgact ctgaccaccc 300
 36 gcagctctct gttcttttga gcccggagta actaggtaac atttaggaac ctccaaaggg 360
 37 tagaagaggg gagtgggtgg gcgtactcta gtcccgcgtg gactgacctc taagtccag 420
 38 actgtcacac ccccttccca ttttttccca acctcaggat ggcgcctcat cccttgatg 480
 39 cgcccaccat ccaagtgtcc caagagaccc agcaaccctt tcccggagcc tcggaccacg 540
 40 aagtgtctag ttccaattcc accccacctc cgtaccgagg gactgctccg 600
 41 aagcagaagc aggtgactgc cgagggacat cgaggaagct ccgtgcgcgg cgcgaggggc 660
 42 gcaacaggcc caagagcgag ttggcactga gcaagcagcg acgaagccgg cgcaagaagg 720
 43 ccaacgaccg ggagcgcaac cgcattgcaca accttaactc cgcgtggat gcgctgcgcg 780
 44 gtgtcctgcc caccttcccg gatgacgcca aacttacaaa gatcgagacc ctgcgttcg 840
 45 ccacaaacta catttgggca ctgactcaga cgctgcgcac agcggaccac agcttctacg 900
 46 gcccagagcc cctgtgccc tgtggggagc tgggaagccc gggagggggc tccagcggcg 960
 47 actggggctc tatctactcc ccagtttccc aagctggtag cctgagcccc acagcctcat 1020
 48 tggaggagtt ccctggcctg caggtgcccc gctccccatc ctgtctgctc ccgggcaccc 1080
 49 tgggtgttct agacttcttg tgaagggccc aaacaggccc tgggcggtgg gcgctggcag 1140
 50 aaagggaggg agtcagagct gtctgaaatg gaaggtagtg gaggcactcg agcatctcgc 1200
 51 cccttctggc ttccattagt caggtccctg atttaaccag gattcgacac gttccttgct 1260
 52 gctgtgcgtg cacaaggac attgcaggct gatctcctct taacctcct cagtgtggcc 1320
 53 acctcaaaact ccgctccaa gcagaggaga gccgtagcac taaatagttg ggagactccc 1380
 54 atacttcctg gtgactccgc cctctttcaa atctgcgggc ctccaaccac cgctttctcc 1440
 55 agagtgcact aatccagtgt
 58 <210> SEQ ID NO: 2
 59 <211> LENGTH: 24
 60 <212> TYPE: DNA

*If you include unknown base residues
 in your nucleotide sequences, you
 must enumerate your unknowns in
 fields 221, 222 and 223.*

RAW SEQUENCE LISTING

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:44

Input Set : A:\ST96042Asq1t.txt

Output Set: N:\CRF3\11082001\I595947A.raw

61 <213> ORGANISM: Artificial Sequence

63 <220> FEATURE:

64 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer

66 <220> FEATURE:

67 <223> OTHER INFORMATION: n = Inosine

69 <400> SEQUENCE: 2

W--> 70 aatkhgm⁶agcgndkcg cryg

73 <210> SEQ ID NO: 3

74 <211> LENGTH: 24

75 <212> TYPE: DNA

76 <213> ORGANISM: Artificial Sequence

78 <220> FEATURE:

79 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers

81 <400> SEQUENCE: 3

82 ggcsrdtytc agggtsybga yctt

85 <210> SEQ ID NO: 4

86 <211> LENGTH: 25

87 <212> TYPE: DNA

88 <213> ORGANISM: Artificial Sequence

90 <220> FEATURE:

91 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers

93 <400> SEQUENCE: 4

94 aaccttaact ccgcgctgga tgcgc

97 <210> SEQ ID NO: 5

98 <211> LENGTH: 18

99 <212> TYPE: DNA

100 <213> ORGANISM: Artificial Sequence ✓

102 <220> FEATURE:

103 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers ✓

105 <400> SEQUENCE: 5

106 cgcggtgtcc tgcccacc

109 <210> SEQ ID NO: 6

110 <211> LENGTH: 6

111 <212> TYPE: DNA

112 <213> ORGANISM: Artificial Sequence ✓

114 <220> FEATURE:

115 <223> OTHER INFORMATION: Description of Artificial Sequence: E box ✓

117 <400> SEQUENCE: 6

118 caggtg

121 <210> SEQ ID NO: 7

122 <211> LENGTH: 6

123 <212> TYPE: DNA

124 <213> ORGANISM: Artificial Sequence ✓

126 <220> FEATURE:

127 <223> OTHER INFORMATION: Description of Artificial Sequence: Mutated E box ✓

129 <400> SEQUENCE: 7

130 tccgtg

133 <210> SEQ ID NO: 8

134 <211> LENGTH: 214

FYI: It is preferred that you describe
your unknown in fields 221 as well as
222 and 223.

RAW SEQUENCE LISTING

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt

Output Set: N:\CRF3\11082001\I595947A.raw

```

135 <212> TYPE: PRT
136 <213> ORGANISM: Rattus norvegicus
138 <400> SEQUENCE: 8
139 Met Ala Pro His Pro Leu Asp Ala Pro Thr Ile Gln Val Ser Gln Glu
140   1           5           10           15
142 Thr Gln Gln Pro Phe Pro Gly Ala Ser Asp His Glu Val Leu Ser Ser
143           20           25           30
145 Asn Ser Thr Pro Pro Ser Pro Thr Leu Val Pro Arg Asp Cys Ser Glu
146           35           40           45
148 Ala Glu Ala Gly Asp Cys Arg Gly Thr Ser Arg Lys Leu Arg Ala Arg
149           50           55           60
151 Arg Gly Gly Arg Asn Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
152   65           70           75           80
154 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
155           85           90           95
157 His Asn Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
158           100          105          110
160 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
161           115          120          125
163 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
164           130          135          140
166 Ser Phe Tyr Gly Pro Glu Pro Pro Val Pro Cys Gly Glu Leu Gly Ser
167 145           150          155          160
169 Pro Gly Gly Gly Ser Ser Gly Asp Trp Gly Ser Ile Tyr Ser Pro Val
170           165          170          175
172 Ser Gln Ala Gly Ser Leu Ser Pro Thr Ala Ser Leu Glu Glu Phe Pro
173           180          185          190
175 Gly Leu Gln Val Pro Ser Ser Pro Ser Cys Leu Leu Pro Gly Thr Leu
176           195          200          205
178 Val Phe Ser Asp Phe Leu
179           210
182 <210> SEQ ID NO: 9
183 <211> LENGTH: 1330
184 <212> TYPE: DNA
185 <213> ORGANISM: Homo sapiens
187 <400> SEQUENCE: 9
188 cctcggaaccc cattctctct tcttttctcc tttggggctg gggcaactcc caggcggggg 60
189 cgctgcagc tcagctgaac ttggcgacca gaagcccgt gagctcccca cgccctcgc 120
190 tgctcatgc tctctattct tttgcgcgg tagaaaggta atatttgag gccttcgagg 180
191 gacgggcagg ggaagaggg atcctctgac ccagcgggg ctgggaggat ggctgtttt 240
192 gttttttccc acctagcctc ggaatcgcg actgcgccg gacggactca aacttacct 300
193 tccctctgac cccgcgtag gatgacgct caaccctcg gtgcgccac tgtccaagt 360
194 acccgtgaga cggagcggc cttccccaga gcctcggaag acgaagtga ctgccccac 420
195 tccgccccgc ccagccccac tcgcacacc gggaactgc cagaggcgga agagggaggc 480
196 tgccgagggg ccccgaggaa gctccgggca cggcgggg gacgcagcc gcctaagagc 540
197 gagttggcac tgagcaagca gcgacggag cggcgaaaga agccaacga ccgcgagcgc 600
198 aatogaatgc acgacctcaa ctcggcactg gacgccctgc gcggtgtct gcccacctc 660
199 ccagacgacg cgaagctcac caagatcgag acgtgcgct tcgcccacaa ctacatctg 720
200 gcgctgactc aaacgctgcg catagcggac cacagcttgt acgcgctgga gccgccggcg 780

```

RAW SEQUENCE LISTING

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt

Output Set: N:\CRF3\11082001\I595947A.raw

```

201 ccgcaactgcg gggagctggg cagcccaggc ggtccccccg gggactgggg gtccctctac 840
202 tccccagctct cccaggctgg cagcctgagt cccgccgcgt cgctggagga gcgaccggg 900
203 ctgctggggg ccacctcttc cgctgcttg agcccaggca gtctggcttt ctcagatttt 960
204 ctgtgaaagg acctgtctgt cgctgggctg tgggtgctaa gggtaaggga gagggaggga 1020
205 gccgggagcc gtagagggtg gccgacggcg gcggccctca aaagcacttg ttccttctgc 1080
206 ttctccctag ctgacccttg gccggcccag gcctccacgg gggcggtagg ctgggttcat 1140
207 tccccggccc tccgagccgc gccaacgcac gcaacccttg ctgctgcccg cgcgaagtgg 1200
208 gcattgcaaa gtgcgctcat tttaggcctc ctctctgcca ccaccccata atcccattca 1260
209 aagaatacta gaatggtagc actaccggc cggagccgc caccgtcttg ggtcgcccta 1320
210 ccctcactca 1330

```

213 <210> SEQ ID NO: 10

214 <211> LENGTH: 214

215 <212> TYPE: PRT

216 <213> ORGANISM: Homo sapiens

218 <400> SEQUENCE: 10

```

219 Met Thr Pro Gln Pro Ser Gly Ala Pro Thr Val Gln Val Thr Arg Glu
220 1 5 10 15
222 Thr Glu Arg Ser Phe Pro Arg Ala Ser Glu Asp Glu Val Thr Cys Pro
223 20 25 30
225 Thr Ser Ala Pro Pro Ser Pro Thr Arg Thr Pro Gly Asn Cys Ala Glu
226 35 40 45
228 Ala Glu Glu Gly Gly Cys Arg Gly Ala Pro Arg Lys Leu Arg Ala Arg
229 50 55 60
231 Arg Gly Gly Arg Ser Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
232 65 70 75 80
234 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
235 85 90 95
237 His Asp Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
238 100 105 110
240 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
241 115 120 125
243 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
244 130 135 140
246 Ser Leu Tyr Ala Leu Glu Pro Pro Ala Pro His Cys Gly Glu Leu Gly
247 145 150 155 160
249 Ser Pro Gly Gly Pro Pro Gly Asp Trp Gly Ser Leu Tyr Ser Pro Val
250 165 170 175
252 Ser Gln Ala Gly Ser Leu Ser Pro Ala Ala Ser Leu Glu Glu Arg Pro
253 180 185 190
255 Gly Leu Leu Gly Ala Thr Ser Ser Ala Cys Leu Ser Pro Gly Ser Leu
256 195 200 205
258 Ala Phe Ser Asp Phe Leu
259 210

```

262 <210> SEQ ID NO: 11

263 <211> LENGTH: 18

264 <212> TYPE: DNA

265 <213> ORGANISM: Artificial Sequence *OK*

267 <220> FEATURE:

268 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer *OK*

RAW SEQUENCE LISTING

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:44

Input Set : A:\ST96042Asq1t.txt

Output Set: N:\CRF3\11082001\I595947A.raw

270 <400> SEQUENCE: 11
 271 caacgaccgg cagcgcaa 18
 274 <210> SEQ ID NO: 12
 275 <211> LENGTH: 24
 276 <212> TYPE: DNA
 277 <213> ORGANISM: Artificial Sequence
 279 <220> FEATURE:
 280 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 282 <400> SEQUENCE: 12
 283 gccagatgt agttgtgggc gaag 24
 286 <210> SEQ ID NO: 13
 287 <211> LENGTH: 60
 288 <212> TYPE: DNA
 289 <213> ORGANISM: Artificial Sequence
 291 <220> FEATURE:
 292 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 294 <400> SEQUENCE: 13
 W--> 295 atcggtgaga ctggtaccag cagagtcacg agagagacta caggtactg gnnnnnnnn 60
 298 <210> SEQ ID NO: 14
 299 <211> LENGTH: 20
 300 <212> TYPE: DNA
 301 <213> ORGANISM: Artificial Sequence
 303 <220> FEATURE:
 304 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 306 <400> SEQUENCE: 14
 307 agacgacgcg aagctcacca 20
 310 <210> SEQ ID NO: 15
 311 <211> LENGTH: 24
 312 <212> TYPE: DNA
 313 <213> ORGANISM: Artificial Sequence
 315 <220> FEATURE:
 316 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 318 <400> SEQUENCE: 15
 319 gctcaccaag atcgagacgc tgcg 24
 322 <210> SEQ ID NO: 16
 323 <211> LENGTH: 25
 324 <212> TYPE: DNA
 325 <213> ORGANISM: Artificial Sequence
 327 <220> FEATURE:
 328 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 330 <400> SEQUENCE: 16
 331 atcggtgaga ctggtaccag cagag 25
 334 <210> SEQ ID NO: 17
 335 <211> LENGTH: 25
 336 <212> TYPE: DNA
 337 <213> ORGANISM: Artificial Sequence
 339 <220> FEATURE:
 340 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
 342 <400> SEQUENCE: 17

No possible values are given for n. No location information is given for n. Therefore Error
Must include fields 221, 222 and 223 responses to explain nucleotide presence.
Good

VERIFICATION SUMMARY

DATE: 11/08/2001

PATENT APPLICATION: US/09/595,947A

TIME: 07:45:45

Input Set : A:\ST96042Asqlt.txt

Output Set: N:\CRF3\11082001\I595947A.raw

L:70 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:2
L:70 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:2
L:70 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:295 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:13
L:295 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13
L:295 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13